

CLAIMS

WHAT IS CLAIMED IS:

1. A speech model training technique for speech recognition, including the
5 following steps:
separating the inputted speech into a compact speech model with clean voice
and an environmental interference model;
filtering out the environmental effects of the inputted speech according to the
environmental interference model and obtaining a speech signal; and
10 plugging the speech signal into the compact speech model and deriving a
speech training model by using the discriminative training algorithm so as to
provide the speech recognition device with the speech training model for
subsequent speech recognition processing.
2. The speech model training technique for speech recognition as claimed in
15 claim 1, wherein the signals of the environmental interference model include
a channel signal and noise.
3. The speech model training technique for speech recognition as claimed in
claim 2, wherein the channel signal includes microphone channel effect.
4. The speech model training technique for speech recognition as claimed in
20 claim 2, wherein the channel signal includes the speaker bias.
5. The speech model training technique for speech recognition as claimed in
claim 1, wherein the discriminative training technique is a generalized
probabilistic descent (GPD) training technique.
6. The speech model training technique for speech recognition as claimed in

claim 1, wherein the step of separating the inputted speech is to compare the non-speech output of the Recurrent Neural Network (RNN) with a predetermined threshold to detect the non-speech frames, and then apply the non-speech frames for calculating the on-line noise model.

5 7. The speech model training technique for speech recognition as claimed in claim 1, wherein the step of filtering out the environmental effects is performing by a filter.

8. The speech model training technique for speech recognition as claimed in claim 1, wherein the step of filtering out the environmental effects further
10 includes the following steps:

employing the state-based Wiener filtering method to process the inputted speech so that the compact speech model can become an enhanced speech;

15 converting the enhanced speech into a Cepstrum Domain to estimate the channel bias by the signal bias compensation (SBR) method and then converting the compact speech model into a bias-compensated speech model; and

employing the parallel model combination (PMC) method and the on-line noise model to convert the bias-compensated speech model
20 into noise- and bias-compensated speech models.

9. The speech model training technique for speech recognition as claimed in claim 8, wherein the signal bias-compensated method is to employ a codebook to encode the feature vectors of the enhanced state-based speech and then calculate the average encoding residuals, wherein the codebook is

formed by collecting the mean vectors of mixture components in the compact speech models.